AN ARCHETYPE PATTERN DRIVEN APPROACH TO ONTOLOGY DEVELOPMENT IN HEALTH CARE INFORMATION SYSTEMS

Wullianallur Raghupathi and Amjad Umar
Graduate School of Business, Fordham University, 113 W. 60th Street, NY, NY 10023
Raghupathi@fordham.edu, Umar@fordham.edu

ABSTRACT

In this exploratory research we examine the potential of archetype patterns for ontology development in health care information systems (HCISs). Archetype patterns are based on the integration of the concepts of archetype and design patterns. The UML provides the underlying modeling support. Our research confirms that the archetype patterns approach provides meta models than can serve as ontologies in health care. Domain ontologies for intra- and inter-enterprise applications can be articulated based on archetype patterns. The patterns, in turn, can provide a foundation for semantic interoperability across health information systems. Future research can focus on governance, standards, and tools for archetype patterns, as well as development of a comprehensive set of health care patterns.

Keywords: **archetype pattern, health care information system (HCIS), ontology**

INTRODUCTION

Today's healthcare delivery systems have experienced a simultaneous increase in costs and decrease in quality [p. 17 in 25]. The result: pressure on healthcare delivery organizations to take drastic and decisive steps [26]. Meanwhile, health care information system (HCIS) designers must focus on the challenge of organizing health information more effectively, thereby enabling participants to cope with these challenges [18]. Needed now are new approaches to modeling high-level architectures across healthcare delivery organizations. And these approaches must dovetail with today's rapid progress in several HCIS areas — among them service-oriented architecture (SOA) and web services in healthcare [9, 20, 27 and 28], open source software development [29, 34], and model-driven architecture (MDA) [4, 30]. Each of these areas is characterized by inter-enterprise complexity and the need for interoperability and standards, which require flexible and robust designs.

Ontologies [16] play a significant role in defining the high-level commonalities in such system requirements. Service definition, for example, is a critical challenge in the use of SOA. An ontology can assist organizations in accepting standard definitions. Likewise, an ontology can lead to more successful open source applications by identifying negotiated design principles and standards. Lastly, an ontology can assist in the description of the platform independent model (PIM).

Several definitions of ontology are found in the literature. An *ontology* for HCIS is concerned "with the principled definition of healthcare classes and the relations among them [adapted from 5 and 6]." An *archetype pattern* is a "high-level abstraction of a generalized solution that can be implemented and applied to solve problems that are common to different health solutions [adapted from 1]." Archetype patterns have the potential to offer a cohesive, integrated framework for ontology development in HCISs. They have the potential to deliver decentralized, modular, loosely coupled and reusable ontologies for HCIS by acting as a gateway to standards and providing an intermediate step in development (e.g. a platform-independent model in MDA).

In this paper, we discuss the application of archetype patterns to modeling ontologies for HCISs. The ontological framework allows high level abstractions and serves as a bridge to specific domain ontologies (e.g. medical/clinical) and interoperability standards. By serving as a frame of reference, the patterns can be used to develop more specific representations of healthcare processes. Using the resultant archetype patterns as meta-level models, healthcare organizations can develop HCISs that comply with universally accepted process definitions and accepted standards.

We follow a brief discussion of ontology development in HISs with an examination of the potential of archetype patterns to model and develop ontologies for HCIS applications. We then offer conclusions and future research directions.

ONTOLOGY DEVELOPMENT IN HCIS

As Bodenreider & Burgun [5] suggest, ontologies "may be characterized according to the domain they represent the level of detail they provide." There are an increasing number of projects underway to develop formal representations of bioinformatics and medical terminology [5] and concepts [for a review of biomedical ontologies, see 5 & 6]. But advances in HCIS ontology development have been limited to exploratory research on archetypes at domain level for electronic health records (EHRs), a sub-set of HCIS [12, 18].

The development of inter-enterprise HCIS applications (e.g. a distributed patient admission system) —using an ontology and reusable patterns of health processes — makes sense in today's era of the SOA-enabled legacy system migration. To move forward with this plan, we consider an ontology (using archetype patterns) that provides an organizing framework. Archetype patterns can assist in SOA governance issues such as "service definition" (e.g. What constitutes a service? Who publishes it?) in an SOA environment. This plan typically would support large integrated digital health systems (e.g. a distributed EHR) in a regional health information network spanning multiple organizations. Thus, an archetype pattern could be a "service definition" or a "publisher description." A critical question is then: How do healthcare delivery organizations retain their unique internal processes and still enable inter-organizational data exchange?

ARCHETYPE PATTERNS

Archetypes were first designed in the Good Electronic Health Record (GEHR) project and its successor, the open EMR project. An archetype can be described as a template for a healthcare (medical) domain concept. A healthcare domain concept (which may vary from the general, such as “health history”, to the detailed, as in “insurance claim status”) comprises basic data structure. For example, a data structure of AddressType consists of a street field, a zip code field, a city field and a state field, and it has an attribute that constrains the address as work address or most recent address. Every piece of healthcare data that is stored in the application is an instance of an archetype [12].

Archetypes and patterns have been discussed in management and organization theory literature for the past several decades. There, archetypes have been described for strategy formulation [23], understanding strategic change [13], organizational transition [24], and much more. Rowland [31] describes archetypes of system design. Schwaninger [32] discusses the use of archetypes to model complexity. In computer science, Massingill [21] and Massingill et al. [22] describe design patterns for parallel programming. Tambouratzis & Tambouratzis [36] identify patterns in neural networks.
Patterns, too, are recognized as a useful way to capture lessons learned and to disseminate and apply practices that have proven successful [17]. According to Gamma et al. [11], a pattern “is a solution to a problem in context.” Typically, a pattern comprises a description of the problem, the context of the problem, and a possible solution to that problem [1]. Archetypes, at a high level of abstraction, “arise from a collective of human experiences (the collective unconscious), that uses archetypes as one of its ordering and structuring principles [1].”

Patterns provide guidance for working top-down from health process design to application runtime. In general, the patterns provide a description of proven practices (and, thus, reusable assets) that identify how to meet a given set of objectives within a specific context. The patterns are generally used in a prescriptive fashion, guiding development from the top down. Archetype patterns, therefore, function at a higher level of abstraction than normal classes. From a conceptual viewpoint, archetype patterns recognize and model universalities while classes are more local and operational. This offers great modeling benefits in healthcare since the high level, abstract archetypes and patterns bridge the gap between internal ontology and universal data messaging standards in HCISs. From an operational perspective, archetype patterns can produce specific classes for implementation.

The approach to ontology development proposed here is rigorous and knowledge based. It is derived from the design principle that knowledge and information levels in HCIS should be separate [2, 37]. In this context, archetype patterns serve several purposes. First, they have the potential to enable users in the healthcare domain to formally express their concepts at the enterprise level (enterprise patterns). Second, they can facilitate individual organizational HCISs to comply with inter-enterprise standards (e.g. service definition). Third, via universally agreed upon archetype patterns of healthcare processes, interoperability is enabled [2]. Finally, the patterns themselves provide operational support in the actual implementation of the HCIS. To wit, archetypes patterns can

characterize groups of data and processes that are specific, highly related and organizationally meaningful to the healthcare delivery enterprise. They have the capability to define the healthcare delivery rules (constraints) for legitimate values, and they use specific terminology (contingent upon a modeling scheme such as the UML) to identify components within an archetype pattern. Using notations and symbols, the patterns can robustly model the processes at a high level of abstraction.

Furthermore, as the archetype patterns are decoupled from the underlying health record model, new archetype patterns can be added over time, allowing a health record system to evolve without substantial changes (similar to archetypes, as reported in [18]). From the ontology that describes the high level inter-enterprise view of the range of processes and data within it (for example, a regional health information network), one can drill down to the domain ontologies, be they clinical terminologies or archetype-driven EHRs. By accessing clinical standards in use —such as HL7 messages, LOINC, the National Health Data Dictionary, and the GP Minimum Data Set— a more effective solution becomes available [3]. According to Bird et al. [3], a common problem with approaches such as HL7, CEN13606, and CorbaMED is that these do not provide a simple long-term view for the standardization of numerous clinical information structures, including clinical (health) tests, notes, care plans, tests for allergies, and medications. Unlike a domain ontology, an archetype pattern offers a higher level of abstraction and emphasizes organizational or enterprise level processes and policies.

Archetype patterns have not seen significant use in health care, although the GEHR is an example of an archetype approach. It uses two-level modeling wherein a generic health record for the health data is first described, and then the structure is constrained by an archetype. In this use, the archetype limits the structure of certain kinds of information [3]. Some have suggested that archetypes facilitate the accommodation by EHRs to change medical and health service delivery practices over time. In [3], the authors argue that clinicians can document electronically content models (e.g. blood pressure, an ECG result or discharge summary), describe clinical workflow processes (e.g. a patient's evaluation), or incorporate clinical or best practice guidelines [3].

The *open*EHR (http://www.openEHR.org) also uses an archetype-based approach [12]. It is defined in comprehensive open specifications for EHR systems based on the results of the European Union's GEMR project of the early 1990s. According to the authors, the *open*EHR archetype has the potential to enable syntactic interoperability and semantic interoperability. Several researchers are exploring the use of the archetype approach in the formal definition of clinical content consistent with developments in current standards. In addition, the archetype approach is closely linked to clinical terminologies, such as SNOMED CT [12]. The key innovation of the *open*EHR architecture is that it uses archetypes to separate record keeping from clinical data collection and thus enables a longitudinal EHR [15, 18].

In the case of the PropeR EHR System (PropeRWeb), a multidisciplinary EHR system for multidisciplinary use in extramural patient care for stroke patients [37], the initial modeling of archetypes focused on the independent definition of healthcare concepts. It

did not take into account possible relationships between archetypes and between the various instances and versions of an archetype.

TOWARDS ARCHETYPE PATTERNS FOR ONTOLOGY DEVELOPMENT IN HCIS

Archetype patterns are a promising approach to providing an ontology for a shared understanding of health processes and data among the healthcare delivery organizations in a given network. To realize archetype patterns, healthcare organizations need to identify existing archetypes patterns, document these, and create templates for continual upgrade [2, 7 and 8]. A significant advantage of archetype pattern-based interoperability is that it gives healthcare organizations and applications the ability to share archetypes and their meta data among them [7]. We define interoperability as the ability to exchange health data among various applications and networks. It is a key and necessary ingredient for the potential success of HIS. But there are other motivating factors for using archetype (design) patterns [1, 10 and 19] in HCISs. Designing a reusable HCIS is difficult. Identifying appropriate objects and abstractions and then building flexible, modular, reliable code for general use in complex healthcare delivery, particularly when dealing with more than one organization, poses a major design challenge. Time-tested frameworks typically do not emerge suddenly. They emerge over time [19]. Also, communicating enterprise "architectural knowledge" to a diverse set of stakeholders can be difficult. This "knowledge" is often abstract and exists in the mind of the architect, rendering it difficult to translate into a computer program. Furthermore, since frameworks are reusable designs (patterns) and not simply code, they are more abstract than the application, which makes documentation challenging. Patterns allow for detailing its purpose, how the pattern should be used, and the design of the framework. For HCISs, archetype patterns offer a number of strengths: providing a common vocabulary (interoperability), explicitly capturing expert knowledge and trade-offs, aiding in the improvement of developer communication, and promoting ease of use.

To summarize, using archetype patterns for HCIS ontology offers numerous advantages including reusability of the patterns (e.g. via a library/repository) within the organization and across organizations. Plus, the patterns can be extended. For example, the concept Senior Citizen Patient can be defined as a sub concept of the concept Patient. In terms of applicability, in an SOA-based HCIS, a consumer of services needs to know the terms or keywords to use when searching for these services. Simple keyword searches cannot accommodate synonyms (patient vs. client), abbreviations (patient ID, pat ID), or foreign languages (patient file, patient table). The solution is to define corresponding relations (e.g. synonym, abbreviation) in the upper-level ontology. Integration in heterogeneous HCISs can take advantage of the framework of a formal ontology. Inter-enterprise applications, such as on an enterprise service bus, require the linking of vast amounts of data, legacy systems and custom (health) applications spread across internal operations, patterns, providers and patients. Archetype patterns can provide a commonality of process definitions. Although the implementations may be local, integration is possible through shared understanding. Archetype patterns can also evolve with the changing

nature of healthcare delivery. Operationally, patterns are visual, and therefore, easily understood. The patterns promote open standards since a healthcare delivery organization must link to other organizations, a system to other systems [35]. An ontology increases the likelihood processes can be exchanged between different systems.

HCISs reflect the healthcare domain in which they operate. One can expect to find archetype patterns in the healthcare domain, in HISs, and in models of those systems. We call this type of archetype pattern a healthcare archetype pattern. A general participant in the healthcare delivery process is typically a *Participant* archetype pattern. It follows that a Patient (or Provider) is a good example of a *Healthcare* archetype pattern. Patient represents an identifiable, addressable unit that may have a legal status (contextual status). Usually this unit represents a person, and one would expect all healthcare systems to have some concept of Patient. Because the notion of archetype pattern is universal, and archetypal entities indeed exist in healthcare systems, the entities themselves can form additional archetypal patterns. For example, the relationships between the archetype patterns *Patient*, *Visit*, *Diagnosis*, and *Prescription* result in processes that are the basis of virtually every healthcare delivery organization (for example, when a Patient makes a Visit, a process to schedule the visit emerges).

CONCLUSIONS AND FUTURE RESEARCH

This exploratory research suggests the use of archetype pattern framework to bridge the enterprise-domain ontology gap by describing health processes at the highest level. The resulting meta model provides a bridge between internal ontology and external standards and interoperability. Participants in an inter-enterprise HCIS, such as an SOA, need to agree on a set of archetype patterns that best characterize their processes. The high level abstractions of the archetype patterns can enable reusability across varied organizations enabling them to retain their internal ontologies while simultaneously building gateways with standards such as HL7. This high-level abstraction describes the basis for several development approaches presently being explored in HIS. These include the MDA approach [4, 30], the open source software application development approaches [29, 34], the service-oriented architecture [9, 28] and web services [20, 27], as well as the recently proposed independent health record banks (IHRB) [33] or health record banking (HRB) system [14].

To enable comprehensive semantic interoperability, the development and maintenance of archetype patterns need to be coordinated and governed globally across healthcare givers [12, 18]. Future research can also focus on more comprehensive development of archetype patterns for healthcare processes as well as additional work in bridges and gateways to domain ontologies, archetype pattern development tools, standards, and governance.

References Available Upon Request